



ASp

la revue du GERAS

15-18 | 1997

Anglais et français de spécialité

Procédure de constitution d'un corpus attesté d'articles de recherche scientifique en vue d'une étude contrastive

Patrick Bachschmidt



Édition électronique

URL : <http://journals.openedition.org/asp/3005>

DOI : 10.4000/asp.3005

ISSN : 2108-6354

Éditeur

Groupe d'étude et de recherche en anglais de spécialité

Édition imprimée

Date de publication : 1 décembre 1997

Pagination : 133-138

ISSN : 1246-8185

Référence électronique

Patrick Bachschmidt, « Procédure de constitution d'un corpus attesté d'articles de recherche scientifique en vue d'une étude contrastive », *ASp* [En ligne], 15-18 | 1997, mis en ligne le 22 mai 2012, consulté le 30 avril 2019. URL : <http://journals.openedition.org/asp/3005> ; DOI : 10.4000/asp.3005

Ce document a été généré automatiquement le 30 avril 2019.

Tous droits réservés

Procédure de constitution d'un corpus attesté d'articles de recherche scientifique en vue d'une étude contrastive

Patrick Bachschmidt

Introduction

- 1 Pour travailler sur une catégorie de discours scientifique, et plus précisément dans le domaine de la mécanique qui nous intéresse, il faut élaborer une procédure de sélection qui permette à la fois de conférer un relatif degré d'homogénéité textuelle qualitative, et préserver une certaine hétérogénéité fondée sur des critères statistiques suffisamment aléatoires pour garantir une représentativité sérieuse de l'échantillonnage.
- 2 Le but d'un échantillonnage en effet est de fournir suffisamment d'informations pour que des inférences concernant les caractéristiques d'une population puissent être effectuées.

1. Ouvrages de référence

- 3 Les articles du corpus en effet, ont été publiés au cours d'une période de six ans¹, par des professeurs des universités cités comme spécialistes de mécanique pour les Francophones et de *mechanical engineering* pour les Anglophones dans des ouvrages de référence. Il s'agit d'une part du Tableau de classement au 31 décembre 1993 du personnel enseignant titulaire du Ministère (français) de l'Enseignement supérieur et de la Recherche, DGA 9, Sciences, section 60 (Mécanique) et, d'autre part, de *The World of Learning* qui donne la liste des universités par ordre alphabétique et par pays.

2. Choix des articles de recherche dans le domaine de la mécanique

- 4 Nous décrivons ici la manière dont nous recherchons un échantillonnage quantitatif représentatif d'articles de recherche au sein d'une population qualitative représentative de professeurs de l'enseignement supérieur spécialistes de mécanique.

3. Description de la procédure

- 5 La population de départ n'est ni plus ni moins la liste des professeurs des universités français recrutés par le canal du Conseil national des universités (CNU), qui est croisée avec celle de l'Annuaire 1993 du CNRS.
- 6 Cet annuaire des unités et des personnels de recherche, département des sciences pour l'ingénieur, comprend 41 pages de 72 noms, soit 2 952 noms en tout. Son rôle est de juger du niveau scientifique du laboratoire (type et qualité des publications, nombre et qualité scientifique des thèses soutenues, qualité des rapports de thèse et des exposés faits par les chercheurs, ainsi que le nombre de brevets déposés et exploités.)
- 7 La population de départ des auteurs anglophones, compte tenu de l'absence d'instances du type de celui du CNU et surtout du CNRS dans les pays de langue anglaise servant de références géographiques dans le choix des corpora, est constituée par les listes des professeurs des universités britanniques, australiennes et américaines, publiées dans *The World of Learning* cité précédemment et qui sont spécialistes de *Mechanical Engineering*.

4. Procédure de recherche

4.1. Application de la loi normale de Gauss

- 8 Des ouvrages de référence cités précédemment, l'on a sélectionné en tout un premier groupe de 147 professeurs britanniques spécialistes de *Mechanical Engineering* et 24 professeurs australiens que nous considérons appartenir à la sphère linguistique britannique puisque la plupart d'entre eux ont fait leurs études au Royaume-Uni, et 382 professeurs américains répondant au critère de professeurs de *Mechanical Engineering*, toutes universités confondues, soit 553 individus répartis pour un tiers en anglophones de la sphère linguistique britannique et deux tiers pour les anglophones de la sphère culturelle et linguistique américaine, en proportion du nombre de professeurs relevé.
- 9 Or, pour chaque population, le nombre de représentants étant supérieur à 120, l'on peut considérer ce nombre comme grand, et cette population suit donc une loi normale gaussienne ; autrement dit, l'ensemble des individus formant la population est réparti symétriquement autour d'une moyenne de la population. L'avantage de la **loi de Gauss** réside précisément dans le fait qu'elle ne dépend d'aucun paramètre extérieur, et que par voie de conséquence, elle ne dépend pas non plus de la moyenne de la population considérée, qui correspond à n'importe quelle population.
- 10 La loi normale de Gauss implique donc également que la population enseignante des professeurs de l'enseignement supérieur spécialistes de mécanique étant supérieure à 120

dans les deux groupes linguistiques considérés, elle est assimilée à l'infini, et nous sommes donc autorisés à choisir une sous-liste représentative de la population par tirage au hasard dans la liste des professeurs de l'enseignement supérieur spécialistes de mécanique.

- 11 Cette liste d'auteurs tirés au sort, se voit conférer conformément à la loi de Gauss, une improbabilité de 5 % de se trouver en dehors de la loi normale, ce qui revient à dire que notre échantillonnage possède une probabilité maximale de se trouver dans les 95 % de la loi normale gaussienne et par conséquent d'être considéré comme correct. Cette probabilité ne peut être égale à 1 car, sinon, elle serait égale à l'ensemble de la population considérée ; mais pour 99,9 %, le tableau de la loi normale indique que le nombre d'échantillons *ne* doit être égal à 30.
- 12 C'est la raison pour laquelle, notre corpus, que nous voulons réellement représentatif de l'ensemble des articles de mécanique publiés en français et en anglais, se compose de 30 articles rédigés en français et 30 autres rédigés en anglais.

4.2. Tirage au hasard des auteurs

- 13 Les professeurs français sont en effet classés par ancienneté dans un échelon donné (« classe exceptionnelle » premier et deuxième échelons, puis « première classe » troisième, deuxième et premier échelons, et « deuxième classe », du sixième au premier échelon), si bien que nous nous trouvons en présence d'une liste chronologique numérotée de 1 à N.
- 14 Puis, à partir d'une table des nombres au hasard², on édite une suite de nombres aléatoires qui permettra un tirage au sort des professeurs représentatifs dans la liste de 1 à N (chaque nombre au hasard correspond au rang du professeur dans la liste).
- 15 Pour les professeurs anglophones, le premier travail du chercheur consiste à dresser la liste de la population des *professors* en *Mechanical Engineering* telle qu'elle est donnée dans *The World of Learning* cité plus haut. Cette liste de population est constituée d'une suite numérotée (par nos soins) des sous-listes de professeurs de la spécialité de chaque université par pays (Royaume-Uni, Australie, États-Unis).
- 16 Ensuite, l'on applique à cette liste la même procédure de tirage au hasard que pour les auteurs francophones de façon à obtenir un sous-ensemble d'échantillons représentatifs de la population.
- 17 Certains cas de rejet d'un tirage doivent néanmoins être envisagés. Pour chaque nom tiré en effet, la validité du tirage est vérifiée par :
- 18 (i) (auteurs francophones) la concordance entre la liste du CNU et celle des laboratoires agréés par le CNRS ou institutions dites prestigieuses.
- 19 (ii) (auteurs britanniques) la confirmation des thèmes de recherche vis-à-vis de la mécanique (*Mechanical Engineering*) dans *Current Research in Britain, Physical Sciences*, qui sert d'ouvrage de référence pour les scientifiques. En d'autres termes, si l'auteur britannique n'est pas cité dans la rubrique *Research in Progress* (pp. 1-546) en tant que chercheur impliqué dans un ou plusieurs projets de recherche, il sera rejeté comme non pertinent en regard du critère qualitatif.
- 20 (iii) (auteurs américains) par la citation dans l'ouvrage *Who's Who in Science in America*, qui tient compte de deux facteurs, d'une part « the position of responsibility held » et d'autre

part de « the level of significant achievement attained in a career of noteworthy activity ».

- 21 L'appartenance des auteurs à l'espace linguistique francophone ou anglophone dès la naissance ou dès la fréquentation du système éducatif élémentaire doit être vérifiée afin d'éviter de traiter des textes linguistiquement atypiques. Les éditions des ouvrages *Who's Who in Science in America* fournissent toutes les indications utiles à cet égard, puisque les rubriques 2, 3 et 7 indiquent respectivement la profession (*occupation*) de l'individu cité, son lieu et sa date de naissance (*vital statistics*) et son cursus scolaire et universitaire (*education*).

4.3. Recherche de publications (articles de recherche)

- 22 La recherche de publications est ensuite effectuée sur la base de données PASCAL disponible dans toutes les bibliothèques universitaires françaises. PASCAL est proposée par l'Institut National de l'Information scientifique et technique (dépendant du CNRS) et constitue sans doute la plus importante base de données couvrant la quasi-totalité des secteurs scientifiques et contenant près de dix millions de références. Pour chaque auteur en effet, l'on obtient une liste d'articles publiés associés. De celle-ci ne seront retenues que les publications répondant aux critères suivants :
- 23 (i) l'appartenance à des listes de journaux et revues de référence (*ULRICH, International Periodicals Directory, 32nd edition*)³, validées par plusieurs professeurs des universités français et *professors* britanniques et américains sur la base de critères qualitatifs fondés à la fois sur l'intérêt et l'originalité de l'information, ainsi que les qualités rédactionnelles imposées par les comités de lecture (*referees*).
- 24 L'on peut ainsi obtenir un ensemble de journaux et revues qualitatifs d'un sous-ensemble quantitatif de professeurs de l'enseignement supérieur spécialistes de mécanique.
- 25 (ii) La concordance entre la langue de publication de l'article de recherche et la langue maternelle de l'auteur francophone à l'exclusion de tout autre car, suite à un rejet de publications dans le cas où l'article n'aura pas été publié dans la langue de l'auteur, il s'avérera nécessaire de procéder à un second, voire un troisième et peut-être un quatrième tirage tel que décrit précédemment pour obtenir les trente articles statistiquement probants par langue.
- 26 Ce type de recherche peut être d'autant plus long qu'en moyenne, PASCAL ne fait état (ce qui est sans doute la vérité), que de trois articles de recherche publiés par tranche de vingt auteurs désignés par la procédure décrite ci-dessus.
- 27 Les soixante articles une fois sélectionnés, il suffit de se les procurer sur support papier et les scanner pour pouvoir en disposer dans la mémoire de son ordinateur et les consulter ou y appliquer des procédures de recherche d'occurrences comme celle prévue par WORD 5.1. ou tout autre concordancier.

5. Difficultés rencontrées

- 28 Outre la collation des soixante articles passés au crible des validations statistiques quantitatives et qualitatives, le problème majeur qui se présente en général est celui des articles d'auteurs francophones rédigés en langue anglaise.

- 29 Il s'avère en effet aujourd'hui que, pour être lus, nombre d'auteurs francophones parmi les professeurs des universités français spécialistes de mécanique, publient en langue anglaise, considérée comme la langue internationale scientifique⁴.
- 30 Il semblerait également qu'à cet égard, la mécanique soit devenue un domaine scientifique dont la diffusion des résultats de recherche se fasse en langue anglaise même dans des revues et des journaux français (comme c'est le cas par exemple du *Journal européen de la mécanique* qui, malgré son nom, est une revue française, où il est rare de trouver plus d'un article de recherche en français sur douze, par livraison).
- 31 En revanche, et paradoxalement, la recherche de très haut niveau dans certains domaines continue de se publier en langue française, en particulier dans les *Comptes-rendus de l'Académie des Sciences* ou *La Revue française de mécanique*. Mais ces cas sont peu communs, non pas que le niveau de la recherche française soit médiocre, mais ils tiennent à des raisons essentiellement financières : la publication d'une revue française de mécanique coûte en effet fort cher.
- 32 Par ailleurs, la grande majorité des auteurs francophones qui se heurtent à de nombreux problèmes rédactionnels auprès des comités de lecture de revues et journaux de langue anglaise souhaiteraient néanmoins publier leurs travaux en français, tout en reconnaissant que la tendance actuelle n'est pas près de s'inverser.
- 33 Il arrive en effet à cet égard que les Mécaniciens soulignent pudiquement, quoique de manière lucide, que ce phénomène est dû à un certain nombre de facteurs autant psycholinguistiques que sociolinguistiques dans la mesure où les chercheurs mécaniciens français souffrent d'une sorte de complexe d'infériorité face à la supériorité reconnue au niveau universitaire et industriel d'une aire différente (c'est le cas de l'Allemagne dans la machine-outil, des États-Unis dans les calculs de structure, qui font en quelque sorte preuve de protectionnisme linguistique par rapport à une zone d'influence) ; le corollaire étant que ces chercheurs auront tendance à publier en langue étrangère. Cette tendance s'inversera donc dans des domaines où la recherche française est en pointe comme c'est le cas en Tribologie mécanique ou en mécanique des nouveaux matériaux à usage technique.
- 34 Cependant, il n'en demeure pas moins que la mécanique actuelle est parcourue de mouvements sous-jacents liés à des phénomènes de suprématie essentiellement économique et qui se traduisent par l'emploi d'une langue plutôt que d'une autre dans la diffusion de l'information scientifique. Un chercheur lira toujours un article de recherche dans une langue qui n'est pas sa langue maternelle, s'il pense ou sait que l'auteur ou les auteurs dudit article appartient ou appartiennent à une aire d'influence jugée meilleure dans le domaine.

BIBLIOGRAPHIE

Current Research in Britain, Physical Sciences. 1994. Harlow : Longman.

Ross, A. 1989. « What content must be a matter of truth conditions ». *The Philosophical Quarterly* 39/156.

Tables statistiques. 1981. Paris : CERESTA.

The World of Learning. 1993. Londres : Europa Publications Limited.

ULRICH, *International Periodicals Directory*, 32nd edition. 1993. New Providence, NJ : R. R. Bowker.

Who's Who in Science in America. 1993. Wilmette, Ill. : Marquis.

NOTES

1. Une période de six ans nous paraît en effet suffisamment compacte pour ne pas provoquer de distorsion conceptuelle car, comme le fait remarquer A. Ross dans *The Philosophical Quarterly*, vol. 39 n° 156 : 271, « *Speakers divided by centuries, may fail to share the necessary concepts* ».
2. *Tables statistiques* éditées par le Centre d'Enseignement et de Recherche de Statistiques Appliquées, (Paris : CERESTA, 1981).
3. Éditions R. R. Bowker, New Providence, NJ, U.S.A.
4. Dans la rubrique « *Notice to authors* » de la revue française *Journal de Physique*, nous pouvons ainsi lire : « *English is highly recommended. French and German are also accepted* ».

RÉSUMÉS

Pour pouvoir se livrer à une étude contrastive à partir d'un corpus qui soit véritablement représentatif de la discipline concernée, il nous paraît nécessaire de se référer à trente articles de recherche par langue étudiée, couvrant une période de six ans. Les auteurs de ces articles auront été tirés au hasard sur une liste officielle de professeurs des universités spécialistes du domaine ; étant donné que cette population est supérieure à cent vingt individus (assimilable à l'infini), dans chaque groupe linguistique, l'on peut appliquer la loi de Gauss dont le tableau indique que le nombre d'échantillons doit être égal à trente. Il suffit ensuite de consulter la base de données PASCAL (couvrant la quasi-totalité des domaines scientifiques) pour sélectionner un article par auteur en fonction de la qualité scientifique du support médiatique, déterminée par un panel de spécialistes du domaine à partir d'une liste de référence de revues et journaux spécialisés. Cependant, cette procédure ne saurait se passer de quelques précautions d'ordre linguistique et soulève quelques problèmes d'ordre culturel.

Carrying out a contrastive study on a corpus assumed to be fully representative of a scientific domain, involves selecting thirty research articles per studied language over a period of six years. The different authors will be randomly drawn on an official document bearing the names of University Professors having an expertise in the field; and as the number of professors goes beyond one hundred and twenty in each linguistic group, the population is said to be representative of the domain so that the table of the Gaussean law can be applied, showing a sampling of thirty articles. Next, access PASCAL data base which covers most scientific fields, so as to select the articles of the corpus (one per author) in relation to the high scientific level

quality of the journals, as assessed by a panel of peers, in a set of reference journals. However, such a selection also requires careful attention to a few linguistic and cultural parameters.

INDEX

Mots-clés : article de recherche, corpus, discours scientifique, échantillonnage, loi de Gauss

Keywords : corpus, Gaussian law of normal distribution, research article, sampling, scientific discourse

AUTEUR

PATRICK BACHSCHMIDT

Patrick Bachschmidt enseigne l'anglais à l'École nationale supérieure des industries textiles (ENSITM) de Mulhouse, Université de Haute-Alsace. Il est titulaire d'un doctorat consacré à la linguistique du discours de la mécanique. p.bachschmidt@uha.fr